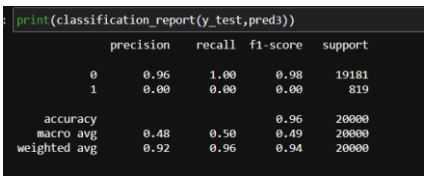
Model Development Phase Template

Date	10 July 2024
Team ID	SWTID1720001058
Project Title	Panic Disorder Detection
Maximum Marks	6 Marks

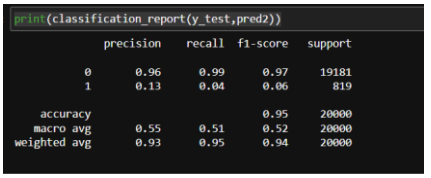
Model Selection Report

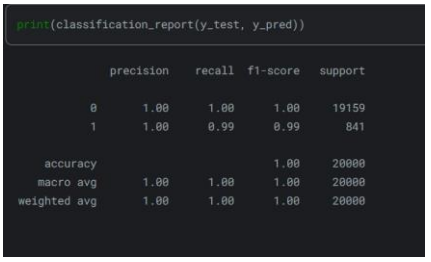
In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

Model Selection Report:

Model	Description	Hyperparameters	Performance Metric (e.g., Accuracy, F1 Score)
Model 1	**AdaBoost (Adaptive Boosting)**: • **Type**: Ensemble method (specifically boosting). • **Description**: AdaBoost works by combining multiple weak classifiers (learners that are slightly better than random guessing, such as decision trees) into a strong classifier. It does this sequentially, adjusting weights on incorrectly classified instances to focus on difficult cases, thereby	• **Number of Estimators**: The number of weak learners to sequentially train. • **Learning Rate**: Controls the contribution of each weak learner to the final combination. Lower rates typically require more estimators for equivalent performance. • **Base Estimator**: The type of weak learner used	

	improving accuracy with each iteration.		

Model 2	• **Gradient Boosting**:
	  • **Type**: Ensemble method (boosting).
	  • **Description**: Similar to AdaBoost, Gradient Boosting sequentially combines weak learners (often decision trees) to create a strong learner. However, unlike AdaBoost which adjusts instance weights, Gradient Boosting fits each new model to the residual errors made by the previous models, gradually reducing the overall error by minimizing a loss function.
		• **Number of Estimators**: Similar to AdaBoost, this specifies the number of boosting stages (weak learners) to be used.
		• **Learning Rate (Shrinkage)**: Controls the contribution of each tree to the ensemble. Lower values require more trees to maintain model complexity.
		• **Tree-Specific Parameters**: Parameters that control the construction of each decision tree, such as maximum depth, minimum samples per leaf, and maximum number of features considered for splitting.
			

Model 3	• **Random Forest**:
	  • **Type**: Ensemble method (bagging).
	  • **Description**: Random Forest builds multiple decision trees during training. Each tree is trained on a random subset of the data and a random subset of the features. During prediction, the ensemble of trees votes to make a final prediction, which helps to reduce
		• **Number of Estimators**: The number of decision trees in the forest.
		• **Max Features**: The maximum number of features considered for splitting a node. It can be specified as a fraction or an integer.
		• **Min Samples Split**: The minimum number of samples required to split an internal node.
			

	overfitting and improve generalization compared to a single decision tree.		
Model 4	**Regression Model**: • **Type**: Supervised learning model. • **Description**: Regression models predict continuous numerical values based on input data. Popular regression models include linear regression (fitting a linear equation to the data), polynomial regression (fitting a polynomial equation), and support vector regression (finding a hyperplane that best fits the data points).	• **Regularization Parameter (Alpha)**: Controls overfitting by penalizing the complexity of the model. • **Degree of Polynomial (for Polynomial Regression)**: Specifies the degree of the polynomial to be used. • **Normalization (for some models)**: Whether to normalize input features before fitting the model.